

Connectieven in de canon

Over de explicitering van relationele coherentie in vereenvoudigde onderwijsteksten

Bachelorscriptie Taalwetenschap

Student: Emma Windau (s1008187)

Begeleider: Wilbert Spooren

6 juni 2022

Radboud Universiteit Nijmegen

Voorwoord

Beste lezer,

In de afgelopen maanden heb ik mijn bachelorscriptie geschreven onder begeleiding van Wilbert Spooren. In februari begon ik met het maken van een lange lijst met opties voor onderwerpen binnen het thema ‘begrijpelijke taal’. Na enig wikken en wegen koos ik voor het onderzoeken van tekstsamenhang op verschillende tekstniveaus. Dit onderwerp kwam voort uit de gedachte dat de samenhang in vereenvoudigde teksten (bv. voor een laaggeletterd publiek) geringer zou kunnen zijn, omdat in deze teksten meer aandacht wordt besteed aan het schrijven van korte zinnen en het maken van simpele constructies.

Later heb ik het onderwerp ‘tekstsamenhang’ verder gespecificeerd, omdat er veel verschillende factoren zijn die invloed kunnen hebben op tekstsamenhang en ik in het korte tijdsbestek van mijn bachelorscriptie slechts een klein onderdeel zou kunnen onderzoeken. Ik koos voor de explicitering van relationele coherentie, ofwel het gebruik van woorden die relaties tussen tekstsegmenten voor de lezer duidelijk maken. Nu, vele weken, gesprekken, aantekeningen, woorden, schema’s, vragen en analyses verder, is mijn scriptie af en kijk ik met plezier terug op het proces.

Op deze plaats wil ik Wilbert bedanken voor zijn goede begeleiding en enthousiasme. Dankzij zijn duidelijke planning is het me gelukt om te blijven schrijven en mijn scriptie bijtijds in te leveren. Tijdens de begeleidingsgesprekken en in zijn feedback op mijn teksten dacht hij met me mee en stelde hij enkele kritische vragen, waardoor ik uiteindelijk tot een beter resultaat ben gekomen.

Tot slot wil ik Barbara bedanken, omdat ze het geduld had om al mijn overwegingen aan te horen en omdat ze me uit mijn scriptiewerk haalde wanneer dat nodig was.

Ik wens u veel leesplezier!

Emma Windau

Inhoudsopgave

Samenvatting	4
1. Inleiding	4
1.1 Adviezen over schrijven in begrijpelijke taal	5
1.2 Cohesie en coherentie	5
1.3 Markering van relationele coherentie in vereenvoudigde teksten	6
1.4 De huidige analyse.....	8
2. Methode	9
2.1 Materiaal.....	9
2.2 Werkwijze	10
2.3 Statistische analyse.....	11
3. Resultaten	11
3.1 Totale dichtheid van connectieven.....	11
3.2 Dichtheden van verschillende typen connectieven	11
3.3 Verdeling van typen connectieven	13
3.4 Kwalitatieve analyse.....	14
4. Discussie	17
4.1 Suggesties voor vervolgonderzoek.....	19
4.2 Conclusie.....	19
Literatuurlijst	20
Bijlagen	23
Bijlage A: Connectievenlijst in T-Scan	23
Bijlage B: Verwijderde citaten.....	25

Samenvatting

In dit onderzoek is gekeken naar de explicitering van relationele coherentie op verschillende tekstniveaus. Hiertoe zijn teksten uit de Canon van Nederland geanalyseerd met T-Scan (Pander Maat et al., 2014). In de Canon van Nederland worden belangrijke gebeurtenissen uit de Nederlandse geschiedenis in 50 teksten behandeld. Die teksten zijn geschreven op 3 niveaus: niveau 1 (kinderen 8-12 jaar en laaggeletterden), niveau 2 (kinderen 12-15 jaar) en niveau 3 (bovenbouw vo en breed publiek). De volgende kenmerken zijn geanalyseerd: (1) de totale dichtheid van connectieven en (2) dichtheid van verschillende typen connectieven (temporeel, opsommend, contrastief, comparatief en causaal). Daarnaast is gekeken naar de verdeling van voorkomens van de verschillende typen connectieven per tekstniveau. Uit de resultaten blijkt dat er tussen tekstniveaus geen verschillen zijn in de totale dichtheid van connectieven. De dichtheid van temporele connectieven bleek op niveau 3 lager dan op niveau 1 en kleiner dan op niveau 2. De dichtheid van contrastieve connectieven bleek op niveau 3 hoger dan op niveau 2. Uit een korte kwalitatieve analyse bleek echter dat de verschillen in het gebruik van temporele en contrastieve connectieven gedeeltelijk verklaard zouden kunnen worden door kleine inhoudelijke verschillen tussen teksten. Er is dus meer (kwalitatief) onderzoek nodig om de huidige resultaten te verifiëren en in context te plaatsen.

Kernwoorden: relationele coherentie, cohesie, connectieven, automatische tekstanalyse, begrijpelijke taal

1. Inleiding

De laatste jaren is er in Nederland steeds meer aandacht gekomen voor begrijpelijke taal. Veel instanties streven naar duidelijke informatiefolders, brieven en webteksten. Zo schrijft de Rijksoverheid op haar website: “Webteksten moeten begrijpelijk zijn voor de bezoekers van een website. Het helpt om teksten te schrijven op een lager taalniveau, zoals B1 en A2. De redactie van Rijksoverheid.nl gebruikt deze taalniveaus” (Dienst Publiek en Communicatie, z.d.). Om instanties te helpen bij het schrijven van teksten in begrijpelijke taal zijn er verschillende initiatieven opgezet die begrijpelijke taal onder de aandacht brengen en adviezen geven over schrijven in begrijpelijke taal. In 2018 werd bijvoorbeeld de Direct Duidelijk Brigade opgericht. Dit is een organisatie die ambtenaren helpt bij het (her)schrijven overheidsteksten in begrijpelijke taal.

Er wordt veel werk verzet op het vlak van begrijpelijke teksten. En dat is ook nodig, want in Nederland is ongeveer 20 procent van de mensen tussen de 25 en 64 jaar lageropgeleid (SCP, Ridder et al. 2020) en is 11.9 procent van de mensen tussen de 16 en 65 jaar laaggeletterd (ECBO & Stichting Lezen en Schrijven, Buisman & Houtkoop 2014). Als teksten te moeilijk zijn, kan dit gedeelte van de bevolking de informatie niet goed tot zich nemen. Een tekst kan op die manier nooit het gewenste effect hebben; als de lezer de tekst

niet begrijpt, zet de tekst de lezer ook niet tot (de juiste) actie aan. Is de tekst op een lager niveau, dan kan de laaggeletterde lezer wél de juiste informatie uit de tekst halen en op basis van deze informatie actie ondernemen.

1.1 Adviezen over schrijven in begrijpelijke taal

Het belang van begrijpelijke teksten is duidelijk. Maar hoe schrijf je eigenlijk een begrijpelijke tekst? Hierover worden vanuit verschillende instanties adviezen gegeven. De Rijksoverheid streeft, zoals hierboven aangegeven, naar schrijven op B1-niveau en A2-niveau. Volgens de Rijksoverheid houdt dit in dat er duidelijke titels en tussenkoppen worden gebruikt, een actieve schrijfstijl wordt gehanteerd, eenvoudige woorden worden gekozen en dat er korte, duidelijke zinnen worden geschreven (Dienst Publiek en Communicatie, z.d.). Een ander voorbeeld is het advies van Stichting Lezen en Schrijven, een organisatie die zich inspant voor het verminderen en voorkomen van laaggeletterdheid en het toegankelijk maken van dienstverlening. Deze stichting geeft in een folder tips over schrijven in eenvoudige taal voor laaggeletterden. Over tekststructuur worden bijvoorbeeld tips gegeven als: “Werk met een titel en tussenkoppen” en “Alinea’s bestaan uit 2 tot 10 zinnen.” Over zinsbouw zegt de stichting onder andere het volgende: “Schrijf actieve zinnen”, “Houd de werkwoorden zoveel mogelijk bij elkaar en dicht bij het onderwerp” en “Voorkom samengestelde zinnen” (Stichting Lezen en Schrijven, 2021).

1.2 Cohesie en coherentie

Een onderwerp dat in veel adviezen over schrijven in begrijpelijke taal niet of nauwelijks expliciet naar voren komt is de samenhang van tekst. In de wetenschappelijke literatuur wordt hiernaar verwezen met de begrippen *cohesie* en *coherentie*. Allereerst moet er onderscheid gemaakt worden tussen deze twee begrippen. Cohesie verwijst naar de oppervlaktestructuur van een tekst: het gebruik van expliciete signalen die delen van de tekst met elkaar verbinden (Noordman & Maes 2000; Crossley & McNamara 2010). Dit zijn bijvoorbeeld connectieven (bv. ‘maar’ en ‘want’) en referentiële uitdrukkingen, zoals woordherhalingen en pronomina. Met het begrip coherentie wordt verwezen naar de verbanden die de lezer tijdens het lezen legt. Hierbij betreft de lezer, naast de signalen in de tekst, ook diens voorkennis (Noordman & Maes 2000; Crossley & McNamara 2010).

In de literatuur wordt vaak onderscheid gemaakt tussen twee typen coherentie: *referentiële coherentie* en *relationele coherentie* (Sanders, 2001). Referentiële coherentie wordt tot stand gebracht door de verwijzingen die in een tekst worden gemaakt om informatie aan referenten te koppelen. Bij relationele coherentie gaat het om de verbanden tussen verschillende tekstsegmenten en de aard en inhoud van die verbanden (Noordman & Maes, 2000).

1.3 Markering van relationele coherentie in vereenvoudigde teksten

In het huidige onderzoek zal worden ingegaan op de markering van relationele coherentie in vereenvoudigde teksten. Een risico van de hierboven reeds aangehaalde adviezen als “Gebruik korte, duidelijke zinnen” is namelijk dat in vereenvoudigde teksten bepaalde tekstelementen weg worden gehaald, waaronder tekstelementen die voor samenhang zorgen. Evers-Vermeul en Van Rooijen (2017) hebben het in deze context bijvoorbeeld over het achterwege laten van signaalwoorden als ‘daarom’ en ‘bovendien’. In een onderzoek van Liu, Kemper en Bovaird (2009) komt het onderwerp ook aan bod. De auteurs schrijven dat het gebruiken van korte woorden en zinnen om Flesch Reading Ease (Flesch, 1948) te verhogen, zou kunnen leiden tot het weglaten van verbanden die uitgedrukt worden in causale en temporele bijzinnen. In een artikel van Graesser en collega’s (2004) wordt het risico van korte zinnen in algemenere zin beschreven: “Textbook writers are known to shorten sentences in basal readers for the purpose of downsizing the grade levels of their texts. The unfortunate liability of shortening sentences is that the texts end up having lower cohesion and coherence” (Graesser et al., 2004, p. 194).

Op theoretische basis valt te verdedigen dat het (juist in vereenvoudigde teksten) onwenselijk is dat de explicitering van samenhang verloren gaat. In de literatuur wordt namelijk uitgegaan van het idee dat de lezer een zogenaamde ‘mentale representatie’ maakt tijdens het lezen. Hierbij wordt de talige informatie in de tekst gecombineerd met de wereldkennis van de lezer om zo tot begrip van tekstrelaties te komen (Noordman & Maes, 2000). Uitgaande van dit idee, zou de aanwezigheid van talige informatie over tekstrelaties kunnen helpen om tot tekstbegrip te komen. Want, zoals Sanders (2001) stelt: “Lezers hebben een tekst immers niet begrepen als ze geen coherentierelaties leggen, maar alleen de afzonderlijke zinnen hebben geïnterpreteerd” (p. 3). Als de lezer de coherentierelatie op basis van wereldkennis al kan leggen, is er wellicht geen explicitering nodig, maar als de lezer niet de juiste kennis heeft, is de lezer waarschijnlijk gebaat bij een expliciete aanduiding van de relatie.

Naar de invloed van explicitering van coherentierelaties op tekstverwerking en -begrip is ook experimenteel onderzoek gedaan. Uit een aantal onderzoeken blijkt dat coherentiemarkeerders een positieve invloed hebben op tekstverwerking (o.a. Sanders 2001; Van Silfhout, Evers-Vermeul, Mak & Sanders 2014; Cain & Nash 2011) en tekstbegrip (o.a. Degand & Sanders 2002; Sanders, Land & Mulder 2007). Echter, het effect lijkt vaak alleen op te gaan voor zwakkere lezers (Van Dooren, Van den Bergh & Evers-Vermeul 2012; Van Silfhout, Evers-Vermeul & Sanders 2014) en lezers met weinig voorkennis (Kamalski, Sanders & Lentz 2008; McNamara & Kintsch 1996).

Ondanks dat de explicitering van relationele coherentie voor de (laagopgeleide, laaggeletterde) lezer waarschijnlijk prettig zou zijn, wordt er in enkele onderzoeken gesteld

dat er een gebrek is aan coherentiemarkeerders in eenvoudige teksten. Land en collega's (2002) analyseerden geschiedenismethodes voor het vmbo en zagen dat veel teksten in deze methodes korte, eenvoudige zinnen bevatten, wat er volgens de auteurs voor zorgt dat de samenhang niet altijd duidelijk is. De auteurs observeerden dat in veel van de teksten weinig structuurexpliciteringen staan. In de teksten waar wel structuurexpliciteringen worden gebruikt, is de samenhang volgens de auteurs soms alsnog afwezig of onduidelijk.

In een later onderzoek van Land, Sanders en Van den Bergh (2008) werden ook studieboekteksten voor het vmbo geanalyseerd, maar deze keer werd dat gedaan voor verschillende niveaus van het vmbo en verschillende vakken. Uit de analyses blijkt dat teksten voor de basisberoepsgerichte leerweg (vmbo-b) minder connectieven bevatten dan teksten voor de hogere niveaus. Tevens schrijven de auteurs dat in de geanalyseerde studieboeken "weinig teksten voorkomen waarin gebruik gemaakt wordt van linguïstische kenmerken die coherentie optimaliseren en de integratie tussen informatie-eenheden bevorderen" (Land, Sanders & Van den Bergh, 2008, p. 76).

Een nuancerend geluid over samenhang in vereenvoudigde teksten komt van Pander Maat en Ditewig (2017). Zij analyseerden de tekstkenmerken van 50 canonteksten over de Nederlandse cultuur en geschiedenis, die aangeboden worden op drie niveaus. Ze schrijven dat in de eenvoudigere teksten meer referentiële verbanden en meer woordherhalingen te vinden zijn. Ze vonden echter, zo schrijven ze, geen duidelijk patroon voor verbindingswoorden. De auteurs rapporteren: "Onze vereenvoudigingen bevatten (...) niet meer verbindingswoorden" (Pander Maat & Ditewig, 2017, p. 252). Daarnaast waren er volgens de auteurs geen "overall effecten" voor temporele en contrastieve¹ connectieven. Wat betreft de causale connectieven valt uit de resultaten op te maken dat teksten op niveau 1 (het eenvoudigste niveau) meer causale connectieven bevatten dan teksten op niveau 2, maar dat teksten op niveau 3 (het hoogste niveau) óók meer causale connectieven bevatten dan teksten op niveau 2. Voor comparatieve connectieven blijkt uit de resultaten van Pander Maat en Ditewig (2017) dat teksten op niveau 3 meer comparatieve connectieven bevatten dan teksten op niveau 2. Dit verschil werd echter niet gevonden tussen niveau 2 en 1.

¹ In Pander Maat en Ditewig (2017) staat: "Overall effecten ontbraken bij adjectiefconcreetheid, werkwoordconcreetheid, en temporele en comparatieve connectieven. Deze zes kenmerken ontbreken daarom in Tabel 1" (p. 249). Echter, in Tabel 1 (p. 250) staan de comparatieve connectieven wel genoemd met bijbehorende resultaten. Omdat uit het artikel niet met zekerheid kon worden afgeleid welke resultaten bij de comparatieve connectieven horen en welke bij de contrastieve, is er contact opgenomen met de auteurs. Maar tot dusver is er nog geen duidelijkheid gekomen. In de huidige beschrijving van het artikel is daarom aangenomen dat de informatie in Tabel 1 van Pander Maat en Ditewig (2017) de juiste informatie is en dat in de tekst op p. 249 dus "temporele en contrastieve connectieven" had moeten staan.

Ook Crossley, Allen en McNamara (2012) nuanceren het idee dat er in eenvoudige teksten een gebrek zou zijn aan expliciete markeerders van coherentierelaties. Zij onderzochten met Coh-Matrix, een automatische tool om tekstkenmerken te analyseren, 300 nieuwsberichten die op drie niveaus (*beginner*, *intermediate* en *advanced*) waren versimpeld. Ze vonden dat teksten op een lager niveau over het algemeen meer cohesieve kenmerken (waaronder ook referentiële kenmerken) hebben. Als het gaat om expliciete markeerders van verbanden, wordt er volgens Crossley, Allen en McNamara (2012) in *beginner*-teksten meer gebruikgemaakt van ontkenningconnectieven (bijvoorbeeld ‘not’ en ‘neither’) en causale connectieven dan in *advanced*-teksten. Echter, het algemene voorkomen van connectieven bleek niet significant te verschillen tussen de niveaus.

Kortom, in de literatuur wordt de suggestie gewekt dat teksten die bewust op een lager niveau geschreven zijn, over het algemeen waarschijnlijk minder connectieven zullen bevatten, maar er zijn studies gedaan waarin geen significant verschil werd gevonden wat betreft de hoeveelheid connectieven in teksten van verschillende niveaus. Er is in ieder geval geen aanleiding om te denken dat vereenvoudigde teksten over het algemeen meer connectieven bevatten, hoewel dit, op basis van de theorie over de mentale representatie van de lezer, misschien juist wenselijk zou zijn.

1.4 De huidige analyse

Om meer zicht te krijgen op het gebruik van coherentiemarkeerders in vereenvoudigde teksten, is het de moeite waard om nogmaals gedetailleerd te kijken naar de verschillen op het gebied van coherentiemarkeerders tussen teksten van verschillende niveaus. De centrale vraag van het huidige onderzoek is dan ook: hoe ziet de expliciete markering van relationele coherentie eruit in teksten met verschillende niveaus van vereenvoudiging? Om dit te onderzoeken is, naar voorbeeld van Pander Maat en Ditewig (2017), gebruikgemaakt van teksten uit de Canon van Nederland (www.canonvannederland.nl).² Deze teksten zijn op drie verschillende niveaus geschreven. Het doel van het huidige onderzoek is om de mogelijke verschillen tussen deze tekstniveaus te achterhalen op het gebied van markering van relationele coherentie.

In het huidige onderzoek zijn de teksten uit de Canon van Nederland geanalyseerd op het voorkomen van markeerders van relationele coherentie en het voorkomen van verschillende typen van deze markeerders (temporeel, opsommend, contrastief, comparatief en causaal). Op basis van de hierboven geschetste ideeën uit de literatuur, valt te verwachten dat eenvoudigere teksten gemiddeld minder markeringen van relationele coherentie zullen bevatten. Maar met het idee van de mentale representatie in het

² Sinds het onderzoek van Pander Maat en Ditewig (2017) is de Canon van Nederland inhoudelijk herzien door de Commissie Herijking Canon van Nederland. Hierbij zijn er tien vensters (onderwerpen) vervangen door andere vensters en zijn alle vensterteksten herschreven.

achterhoofd, is het voor de hand liggend om te denken dat eenvoudige teksten juist meer markerings bevatten, omdat de lezer voor wie de eenvoudige tekst bedoeld is, minder voorkennis heeft en dus meer baat heeft bij expliciete markerings. Echter, de praktijk kan anders zijn; het is goed mogelijk dat de auteur van de tekst dit idee niet heeft meegenomen bij het schrijfproces.

Over het voorkomen van verschillende typen markeerders valt in ieder geval, op basis van Crossley, Allen en McNamara (2012), te zeggen dat er mogelijk meer causale connectieven voorkomen in eenvoudigere teksten. Op basis van Pander Maat en Ditewig (2017) is de verwachting dat er geen significante verschillen zullen zijn tussen tekstniveaus wat betreft de temporele en contrastieve connectieven. Voor de comparatieve connectieven valt, op basis van Pander Maat en Ditewig (2017), te verwachten dat op niveau 3 meer connectieven van dit type voorkomen dan op niveau 2. Over opsommende connectieven is op basis van de aangehaalde literatuur geen verwachting te formuleren.

2. Methode

2.1 Materiaal

In het huidige onderzoek zijn 50 teksten uit de Canon van Nederland (Canon van Nederland, z.d.)³ geanalyseerd op de markering van relationele coherentie. De teksten in de canon gaan over culturele en historische onderwerpen die belangrijk zijn voor de ontwikkeling van Nederland. De Canon van Nederland werd in 2006 gelanceerd door de Commissie-Van Oostrom en werd in 2020 herzien door de Commissie Herijking Canon van Nederland (Commissie Herijking Canon van Nederland, 2020).

Om de canon breed toegankelijk te maken zijn de 50 teksten op de website in drie versies beschikbaar. Die versies worden als volgt onderscheiden door de Commissie Herijking Canon van Nederland (2020).

- Niveau 1: teksten op B1-niveau, geschikt voor kinderen van 8-12 jaar en voor laaggeletterden.
- Niveau 2: teksten op B2-niveau, geschikt voor jongeren van 12-15 jaar.
- Niveau 3: teksten op C1-niveau, geschikt voor de bovenbouw van het voortgezet onderwijs en voor een breed publiek.

In totaal zijn er dus 150 teksten meegenomen in de analyse; van iedere tekst drie versies.

³ De teksten die het materiaal vormen voor het huidige onderzoek, zijn gekopieerd van www.canonvannederland.nl tussen 5-4-2022 en 17-4-2022. De teksten zijn ook te vinden in het volgende rapport: Commissie Herijking Canon van Nederland (2020). *Open vensters voor onze tijd: de Canon van Nederland herijkt*. Stichting entoenu. Geraadpleegd op 17 maart 2022, van <https://www.canonvannederland.nl/nl/downloads>.

2.2 Werkwijze

De analyse van de teksten is uitgevoerd met behulp T-Scan (Pander Maat et al., 2014), een programma waarmee veel uiteenlopende tekstkenmerken kunnen worden geanalyseerd. T-Scan kan de dichtheid van (verschillende typen) connectieven in een tekst berekenen. Hierbij wordt het begrip ‘connectieven’ vrij breed opgevat als “woorden die veelal complete zinnen in een betekenisrelatie met elkaar plaatsen” (Pander Maat, Kraf & Dekker, 2020, p. 43). Het gaat hier voornamelijk om voegwoorden (bv. ‘omdat’) en voornaamwoordelijke bijwoorden (bv. ‘daarvoor’) en in sommige gevallen tijdbijwoorden (bv. ‘toen’) (Pander Maat, Kraf & Dekker, 2020, p. 43). In Bijlage A staat de lijst met connectieven die door T-Scan meegeteld worden als connectief (Pander Maat, Kraf & Dekker, 2020, p. 71-72).

Voordat de teksten met T-Scan geanalyseerd konden worden, zijn er een aantal bewerkingen uitgevoerd op de teksten. Ten eerste zijn de tussenkopjes in alle teksten verborgen, zodat ze niet meegenomen werden in de analyse. De kopjes zouden anders namelijk wel in het woordenaantal meegenomen worden, terwijl ze geen hele zinnen bevatten en geen onderdeel zijn van de lopende tekst. Ten tweede zijn er enkele typ- en spelfouten verbeterd. Ten derde zijn (lange) citaten met eventueel bijbehorende vertaling (zie Bijlage B) uit de teksten gehaald, omdat citaten per definitie geen zinnen bevatten die door de auteur geschreven zijn. De auteur heeft weliswaar bedacht of het citaat geschikt is voor het tekstniveau, maar binnen de citaten heeft de auteur geen keuzes gemaakt over de opbouw en samenhang van zinnen. Tot slot zijn er nog enkele bewerkingen gedaan om de teksten T-Scan-klaar te maken. Alle kromme aanhalingstekens zijn vervangen door rechte aanhalingstekens, tussen alinea’s is telkens een witregel ingevoegd (ook op plekken waar enkel één enter was gebruikt) en de teksten zijn omgezet naar een txt-formaat in de codering UTF-8.

Alle teksten zijn door T-Scan geanalyseerd op de volgende kenmerken: de totale dichtheid⁴ van connectieven⁵, de dichtheid van temporele connectieven, de dichtheid van opsommende connectieven voor (deel)zinnen⁶, de dichtheid van contrastieve connectieven, de dichtheid van comparatieve connectieven en de dichtheid van causale connectieven.

⁴ In T-Scan is de dichtheid het aantal gevonden gevallen per 1000 woorden.

⁵ Voor de berekening van de totale dichtheid van connectieven zijn de dichtheden van de vijf verschillende typen connectieven per tekst bij elkaar opgeteld.

⁶ De dichtheid van opsommende connectieven voor woordgroepen is niet meegenomen, omdat dit type connectieven vooral verbanden binnen zinnen aanduidt, terwijl de andere typen vooral verbanden tussen (deel)zinnen aanduiden.

2.3 Statistische analyse

De statistische analyse, die ingezet werd om de teksten van verschillende tekstniveaus met elkaar te vergelijken, bestond uit meerdere onderdelen. Allereerst werden de teksten van drie tekstniveaus middels een ANOVA voor herhaalde metingen vergeleken op de totale dichtheid van connectieven. Vervolgens werd dezelfde toets gebruikt om de teksten te vergelijken op de dichtheid van verschillende typen connectieven (temporeel, opsommend, contrastief, comparatief en causaal). In totaal werden er dus zes ANOVA's voor herhaalde metingen uitgevoerd. Tot slot werd er een chi-kwadraattoets uitgevoerd om de verdelingen van voorkomens van typen connectieven op de drie tekstniveaus met elkaar te vergelijken. Ten behoeve van deze chi-kwadraattoets werden, op basis van de dichtheden van de verschillende typen connectieven en het aantal woorden per tekstniveau, de absolute aantallen connectieven per tekstniveau berekend.

3. Resultaten

De vraag die in dit onderzoek centraal staat, is: hoe ziet de expliciete markering van relationele coherentie eruit in teksten met verschillende niveaus van vereenvoudiging? Deze vraag zal in deze paragraaf behandeld worden door te rapporteren over twee kenmerken van de teksten: de dichtheid van (verschillende typen) connectieven (zie paragraaf 3.1 en 3.2) en de verdeling van voorkomens van verschillende typen connectieven (zie paragraaf 3.3). Voor beide kenmerken is de samenhang met de variabele 'tekstniveau' statistisch geanalyseerd. In paragraaf 3.4 zal verslag worden gedaan van een korte kwalitatieve analyse die is uitgevoerd naar aanleiding van de resultaten van de statistische analyses.

3.1 Totale dichtheid van connectieven

Uit een eenwegs ANOVA voor herhaalde metingen met als binnen-subjectfactor 'tekstniveau' (niveau 1, niveau 2, niveau 3) en als afhankelijke variabele 'dichtheid van connectieven' bleek dat er geen significant hoofdeffect was van het tekstniveau op de dichtheid van connectieven, $F(2, 98) = 2.248, p = .111, \eta^2_p = .044$.

3.2 Dichtheden van verschillende typen connectieven

In Tabel 1 zijn de resultaten van vijf eenwegs ANOVA's te zien met als binnen-subjectfactor het tekstniveau (niveau 1, niveau 2, niveau 3) en als afhankelijke variabele de dichtheid van het betreffende type connectief. Mauchly's toetsen toonden aan dat de assumptie van sfericiteit geschonden werd voor de typen 'temporeel' ($\chi^2(2) = 7.466, p = .024$), 'opsommend' ($\chi^2(2) = 12.604, p = .002$), 'comparatief' ($\chi^2(2) = 15.687, p < .001$) en 'causaal' ($\chi^2(2) = 11.278, p = .004$). Om die reden werden bij de ANOVA's voor temporele ($\epsilon = .874$), opsommende ($\epsilon = .812$), comparatieve ($\epsilon = .782$) en causale ($\epsilon = .827$) connectieven Huynh-Feldt-aanpassingen uitgevoerd.

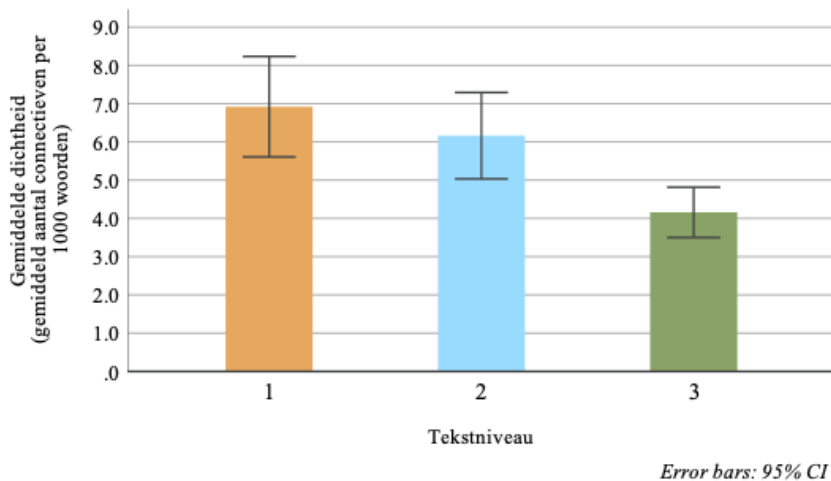
Tabel 1. Resultaten uit ANOVA's voor de dichtheden van verschillende typen connectieven

Type connectief	<i>df</i>	<i>F</i>	<i>p</i>	η^2_p
Temporeel	1.81, 88.57	10.432	< .001	.176
Opsommend	1.67, 81.96	2.998	.064	.058
Contrastief	2, 98	3.840	.025	.073
Comparatief	1.61, 78.71	.964	.369	.019
Causaal	1.70, 83.50	.127	.849	.003

Uit de resultaten (zie Tabel 1) blijkt dat er voor temporele en contrastieve connectieven significante verschillen in dichtheid bestaan tussen tekstniveaus. Voor de andere typen connectieven werden geen significante verschillen gevonden.

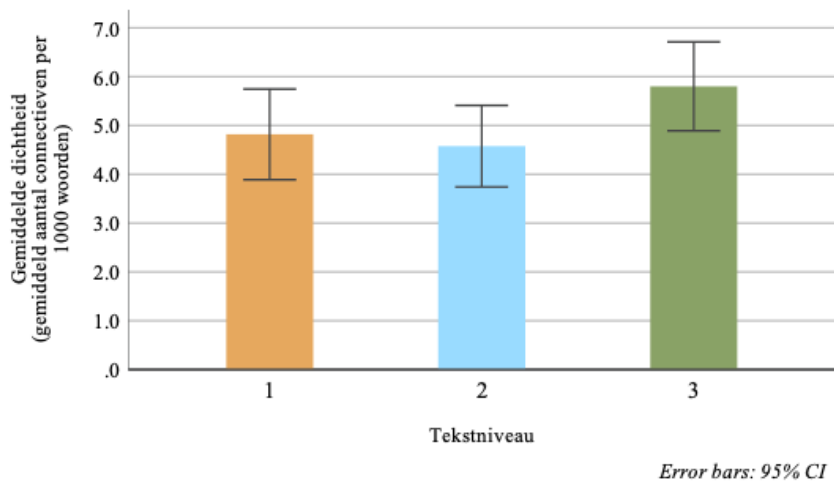
Uit post-hoc-analyses voor temporele connectieven bleek dat de gemiddelde dichtheid van temporele connectieven op niveau 1 significant hoger is dan op niveau 3 (MD = 2.763, $p = .001$). Daarnaast bleek dat de gemiddelde dichtheid van temporele connectieven op niveau 2 ook significant hoger is dan op niveau 3 (MD = 2.004, $p = .006$). Er werd echter geen significant verschil gevonden tussen niveau 1 en 2 ($p = .446$). Figuur 1 is een weergave van de gemiddelde dichtheden van temporele connectieven op de drie verschillende tekstniveaus.

Figuur 1. De dichtheid van temporele connectieven per tekstniveau



Uit post-hoc-analyses voor de categorie van contrastieve connectieven bleek dat de gemiddelde dichtheid op niveau 3 significant hoger is dan op niveau 2 (MD = 1.224, $p = .034$). Er werd geen significant verschil gevonden tussen niveau 1 en 3 ($p = .146$) en ook niet tussen niveau 1 en 2 ($p = 1.000$). Figuur 2 is een weergave van de gemiddelde dichtheden van contrastieve connectieven op de drie verschillende niveaus.

Figuur 2. De dichtheid van contrastieve connectieven per tekstniveau



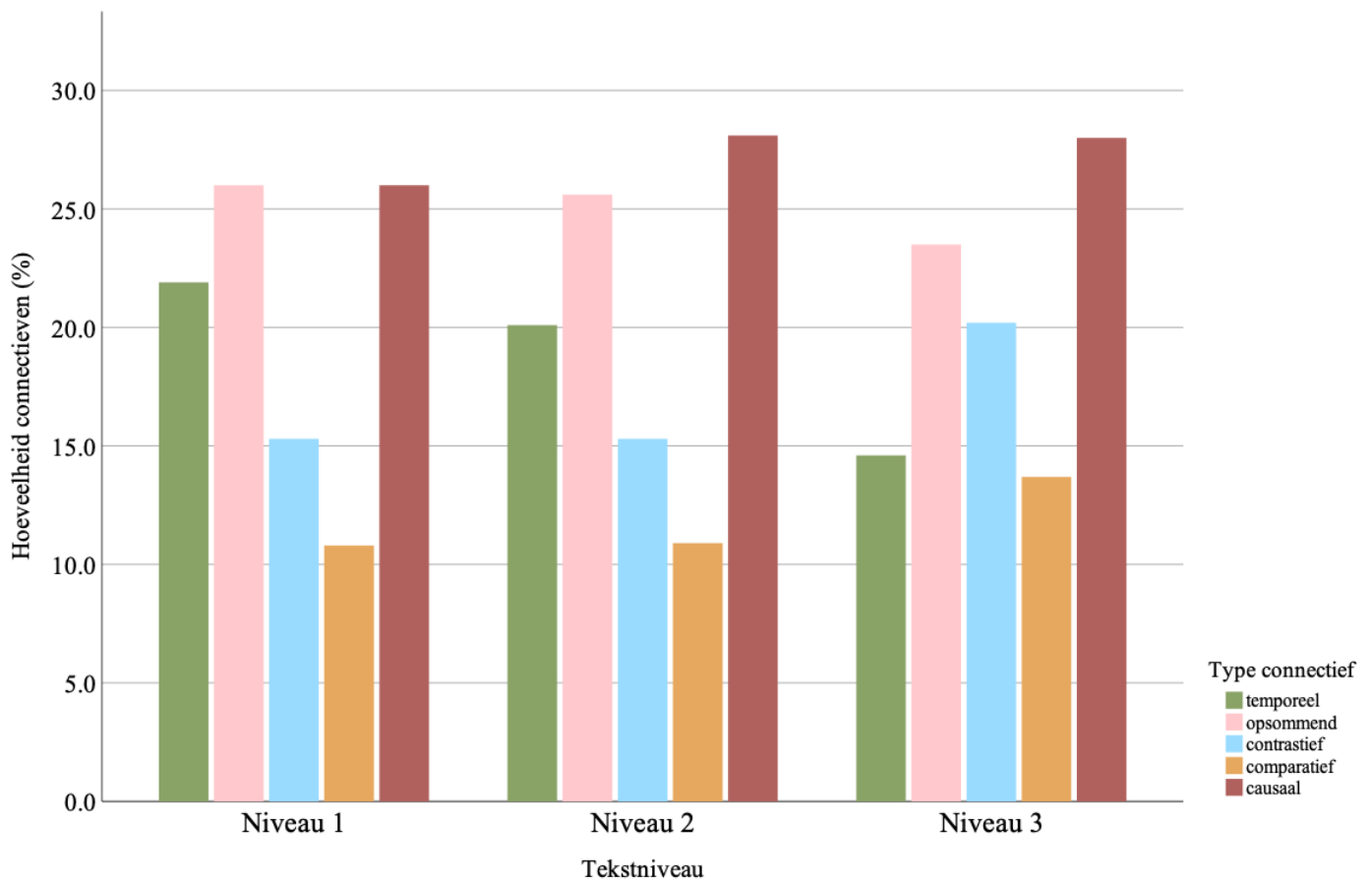
3.3 Verdeling van typen connectieven

Uit een chi-kwadraattoets, gebaseerd op absolute aantallen van de verschillende typen connectieven per tekstniveau, bleek dat er sprake is van samenhang tussen het tekstniveau en de verdeling van voorkomens van de verschillende typen connectieven ($\chi^2(8) = 22.393$, $p = .004$). In Figuur 3 is weergegeven hoe de verdeling van ieder niveau er, uitgedrukt in percentages, uitziet. In Tabel 2 is aan de gestandaardiseerde residuen (tussen haakjes) te zien dat het effect van het tekstniveau bij zowel de temporele als de contrastieve connectieven vooral veroorzaakt wordt door een afwijking op niveau 3. Een belangrijke kanttekening is dat de effectgrootte zeer gering bleek ($V = .075$). Dit duidt erop dat het effect psychologisch verwaarloosbaar klein is.

Tabel 2. Absolute aantallen (en gestandaardiseerde residuen) van typen connectieven per tekstniveau

	Temporeel	Opsommend	Contrastief	Comparatief	Causaal	Totaal
Niveau 1	122 (1.9)	145 (.5)	85 (-1.1)	60 (-.8)	145 (-.7)	557
Niveau 2	137 (1.0)	174 (.4)	104 (-1.2)	74 (-.8)	191 (.3)	680
Niveau 3	114 (-2.5)	183 (-.8)	157 (2.0)	107 (1.4)	218 (.3)	779

Figuur 3. Verdeling van voorkomens van verschillende typen connectieven per tekstniveau



3.4 Kwalitatieve analyse

Om de gevonden verschillen tussen tekstniveaus beter te kunnen begrijpen, is van enkele teksten die grote verschillen voor de relevante typen connectieven vertoonden, kort bekeken welke kwalitatieve verschillen er te zien zijn. Voor de temporele connectieven is tekst 41 (De televisie) op niveau 1 en 3 bekeken en tekst 47 (Het Caribisch gebied) op niveau 2 en 3. Voor de contrastieve connectieven zijn de teksten 30 (De Grondwet) en 49 (Europa) op niveau 2 en 3 bekeken. Er is telkens bekeken hoe de coherentierelaties achter de connectieven op het niveau met een grotere hoeveelheid connectieven, zijn vormgegeven op het niveau met een kleinere hoeveelheid connectieven. Zo is in tekst 41 bekeken hoe de coherentierelaties achter de temporele connectieven van niveau 1 worden vormgegeven op niveau 3.

In de teksten 41 en 47 was er een verschil in het gebruik van temporele connectieven in de volgende situaties.

(a) De tekstinhoud die op niveau 1 of 2 staat, is op niveau 3 weggelaten.

Voorbeeld 1 (uit tekst 41)

Niveau 1: “*Voordat* er televisie was zaten mensen rond de eettafel om samen spelletjes te doen of naar de radio te luisteren.”

Niveau 3: (geen vergelijkbare inhoud gevonden)

(b) Op niveau 3 is voor een andere formulering gekozen; het connectief is vervangen door (een) ander(e) woord(en) (zie Voorbeeld 2a) of is volledig weggelaten (zie Voorbeeld 2b).

Voorbeeld 2a (uit tekst 41)

Niveau 1: “Lang niet iedereen had *meteen* een televisie.”

Niveau 3: “Er zijn in de beginperiode dan ook maar zo'n vijfhonderd Nederlanders die een tv in huis hebben.”

Voorbeeld 2b (uit tekst 47)

Niveau 2: “De eilandbewoners wilden niet langer dat Nederlandse bedrijven een grote invloed hadden op de economie, terwijl *tegelijktijd* veel mensen in armoede leefden.”

Niveau 3: “Inwoners zijn ontevreden over het overwicht van Nederlandse bedrijven op de Curaçaose economie terwijl op het eiland veel mensen in armoede leven.”

In de teksten 30 en 49 was er een verschil in het gebruik van contrastieve connectieven in de volgende situaties.

(a) De tekstinhoud die op niveau 3 staat, is op niveau 2 weggelaten.

Voorbeeld 3 (uit tekst 30)

Niveau 3: “Het gaat hier niet om rechten die de burgers onderling – ten opzichte van elkaar – hebben, *maar* om de rechten die burgers de vrijheid geven om hun leven in te richten zonder dat de staat zich met hun opvattingen en levenskeuzes bemoeit.”

Niveau 2: “Het gaat dan ook over de belangrijkste rechten in een land: rechten die burgers de vrijheid geven om hun leven in te richten zoals ze zelf willen.”

(b) Op niveau 2 is voor een andere formulering gekozen; het connectief is vervangen door (een) ander(e) woord(en).

Voorbeeld 4a (uit tekst 49)

Niveau 3: “De Europese samenwerking wordt steeds intensiever en zal in de jaren die volgen niet alleen over economie gaan, *maar* ook over cultuur, onderwijs en milieu.”

Niveau 2: “Het ging niet meer alleen over handel. Het ging nu ook over onderwerpen als cultuur, onderwijs en milieu.”

Voorbeeld 4b (uit tekst 49)

Niveau 3: “*Ondanks* deze ontwikkelingen is in de meeste van de overgebleven 27 lidstaten de steun voor de EU groot, ook in Nederland.”

Niveau 2: “Toch vinden de meeste mensen in Europa de EU belangrijk. Ze accepteren dat samenwerken met 27 landen, 27 culturen en 27 meningen niet altijd gemakkelijk is.”

In de teksten valt dus te zien dat de verschillen tussen tekstniveaus soms verklaard kunnen worden door weglating (of toevoeging) van inhoud. In die gevallen heeft het wel of niet gebruiken van een connectief niets te maken met het wel of niet expliciteren van een coherentierelatie. Op het moment dat er in de ene tekst voor een andere formulering is gekozen dan in de andere tekst, is er soms wél sprake van verschil in explicitering. Er worden echter ook vaak andere woorden (niet-connectieven of connectieven die niet door T-Scan worden gesignaleerd als connectief) gebruikt om het verband toch duidelijk te maken, zoals in Voorbeeld 4a (niveau 2) gebeurt met de woorden ‘meer’ en ‘nu’. De vraag is dan of er nog echt sprake is van verlies van explicitering van de coherentierelatie.

Bij het lezen van de teksten valt op dat sommige woorden wellicht onterecht door T-Scan worden genegeerd. In Voorbeeld 4b (niveau 2) is dit te zien: ‘toch’ verbindt hier wel degelijk tekstgedeeltes aan elkaar, maar het woord wordt niet geteld als (contrastief) connectief. Dat is begrijpelijk, want het woord ‘toch’ kan vele betekenissen en functies hebben en T-Scan kan niet achterhalen in welke betekenis en functie ‘toch’ hier heeft. Het blijft echter een connectief en het zou dus meegeteld moeten worden. Wellicht houdt de kracht van de computer hier op en moet er uiteindelijk toch een mens aan te pas komen om dit soort gevallen mee te nemen in de analyse.

Een andere observatie is dat het soms lijkt alsof de woorden die T-Scan meeneemt niet de belangrijkste dragers van relationele coherentie zijn. In tekst 41 (De televisie) staat bijvoorbeeld de volgende zin: “Dit programma, gepresenteerd door Mies Bouwman, wordt in 1962 tegelijk op radio en televisie uitgezonden en duurt 23 uur.” Het woord ‘tegelijk’

wordt in dit geval door T-Scan gezien als (temporeel) connectief, terwijl het eigenlijk helemaal niet bijdraagt aan de tekstuele samenhang.

4. Discussie

De resultaten van het huidige onderzoek geven een beeld van de explicitering van relationele coherentie op verschillende tekstniveaus. Er zijn weinig (kwantitatieve) verschillen gevonden tussen tekstniveaus. Allereerst bleek de totale dichtheid van connectieven niet afhankelijk van het tekstniveau. Deze bevinding komt overeen met de bevinding van Crossley, Allen & McNamara (2012), die een onderzoek uitvoerden dat qua opzet gelijkenis vertoont met het huidige onderzoek. Het lijkt er dus op dat het vereenvoudigen van teksten geen invloed heeft op de mate waarin coherentierelaties geëxpliciteerd worden; gemiddeld genomen bevatten eenvoudige teksten niet meer of minder connectieven dan moeilijkere teksten. De angst dat connectieven in eenvoudige teksten weggelaten worden (Evers-Vermeul & Van Rooijen 2017; Liu, Kemper en Bovaird 2009; Graesser et al. 2004), lijkt dus niet uit te komen. Maar er lijken vooralsnog ook niet méér connectieven te worden gebruikt in eenvoudigere teksten.

Wat betreft de dichtheden van verschillende typen connectieven bleek dat er alleen verschillen bestaan tussen tekstniveaus als het gaat om temporele en contrastieve connectieven. Temporele connectieven bleken in teksten van het hoogste niveau (niveau 3) minder te worden ingezet dan in teksten van lagere niveaus.⁷ En contrastieve connectieven bleken in teksten van het niveau 3 meer ingezet te worden dan in teksten van niveau 2. Uit de chi-kwadraattoets bleek daarnaast dat niveau 3 een onverwachte verdeling van typen connectieven heeft. Het afwijken van deze verdeling wordt met name veroorzaakt door een in verhouding groot aantal contrastieve connectieven en een in verhouding klein aantal temporele connectieven op niveau 3.

De huidige resultaten voor de dichtheden van verschillende typen connectieven komen niet geheel overeen met de resultaten van Pander Maat en Ditewig (2017). Zij vonden namelijk geen significante verschillen voor temporele connectieven en contrastieve connectieven. Voor causale en comparatieve connectieven waren er in hun studie wel significante verschillen (zie paragraaf 1), maar de auteurs trekken de conclusie dat er geen heldere

⁷ Bij de kwalitatieve analyse viel het op dat T-Scan het woord ‘eerst’, dat in tekst 30, 41 en 47 voorkwam, niet heeft meegenomen als temporeel connectief. Het woord ‘eerst’ staat echter wel als temporeel connectief in de lijst van T-Scan (zie Bijlage A). In de handleiding van T-Scan (Pander Maat, Kraf & Dekker, 2020) is niet terug te vinden waarom T-Scan het woord ‘eerst’ niet meeneemt. Voor zover te zien in de genoemde teksten, doet T-Scan dit op alle niveaus waarop het woord voorkomt en niet selectief op bepaalde niveaus. Mocht T-Scan dus in alle teksten het woord ‘eerst’ overgeslagen hebben bij de analyse, dan zou dit niet al te veel effect moeten hebben op de resultaten, tenzij ‘eerst’ in verhouding heel veel gebruikt is op niveau 3.

patronen gevonden zijn voor verbindingswoorden. Dat de resultaten van het huidige onderzoek afwijken van die van Pander Maat en Ditewig (2017) zou natuurlijk te maken kunnen hebben met het feit dat de Canon van Nederland in 2020 herijkt is. Bij de herijking zijn de niveaus anders ingedeeld⁸, zijn er herschrijvingen gedaan en zijn er tien onderwerpen vervangen door andere onderwerpen.

De kwalitatieve analyse (zie paragraaf 3.4) geeft iets meer inzicht in de resultaten voor temporele en contrastieve connectieven. De waarneming dat sommige connectieven verdwijnen vanwege de weglating van inhoudelijke elementen, nuanceert de kwantitatieve resultaten. Als het verschil in inhoud geen coherentierelatie betreft, maar het volledige tekstgedeelte waarin de coherentierelatie voorkomt (zoals in Voorbeeld 1), dan is er duidelijk geen sprake van en gebrek aan explicitering van de coherentierelatie. Als er een coherentierelatie wordt weggelaten en daarmee een connectief (zoals in Voorbeeld 3), is er ook geen sprake van gebrek aan explicitering, maar komt wel de vraag naar boven of de auteur wellicht bewust een coherentierelatie heeft toegevoegd of weggelaten, bijvoorbeeld omdat de auteur de coherentierelatie op zichzelf als te moeilijk heeft beoordeeld voor een bepaald niveau.

Verschillen in formulering hebben hier en daar ook gezorgd voor het wegvallen van connectieven. Een verschil in formulering kan daadwerkelijk een verschil in explicitering van de coherentierelatie teweegbrengen, maar mogelijk is dit niet altijd het geval. Bijvoorbeeld wanneer de herformulering woorden bevat die door T-Scan niet gezien worden als connectief, maar wel een connectief zijn (zoals in Voorbeeld 4b). Of als er niet-connectieven gebruikt worden die het verband (bijna) even duidelijk aangeven als het connectief dat zij vervangen. In het geval van temporele coherentierelaties lijkt dat goed mogelijk, omdat de lezer er in principe al vanuit kan gaan dat dat een verhaal op volgorde staat en de volgorde vaak ook kan afleiden uit betekenissen van werkwoorden zoals ‘blijven’ en ‘worden’, uit werkwoordstijden en uit specifieke tijdsaanduidingen. Echter, het blijft de vraag of daarmee in dezelfde mate wordt geëxpliciteerd wat de relatie is.

Kortom, er zijn vanuit de kwalitatieve analyse een aantal kanttekeningen te plaatsen bij de resultaten van de kwantitatieve analyse. Inhoudelijke verschillen tussen teksten, connectieven die niet door T-Scan worden meegenomen en niet-connectieven die ook coherentierelaties duidelijk kunnen maken (al zij het misschien niet in de mate waarin een connectief dat kan) zijn contextuele factoren die meespelen bij de verschillen in connectiefgebruik tussen teksten. Een verschil in connectiefgebruik betekent dus niet automatisch een verschil in explicitering van coherentierelaties.

⁸ Eerder was de indeling als volgt: niveau 1 voor groep 5/6 po en laaggeletterden, niveau 2 voor groep 7/8 en onderbouw vo, niveau 3 voor docenten en geletterden (Pander Maat & Ditewig, 2017).

4.1 Suggesties voor vervolgonderzoek

Zoals blijkt uit bovenstaande overdenkingen en nuanceringen naar aanleiding van de kwalitatieve analyse, kan het nuttig zijn om de teksten (na automatische analyse) door een mens te laten bekijken. T-Scan is goed inzetbaar voor een onderzoek als het huidige, maar aanvullend kwalitatief onderzoek is nodig. In vervolgonderzoek zou een combinatie van kwantitatief en kwalitatief onderzoek het beste beeld opleveren van het gebruik van connectieven op verschillende tekstniveaus.

In wezen is het huidige onderzoek een onderzoek naar samenhang van vereenvoudigde teksten. Echter, slechts een klein deel van de elementen die voor samenhang zorgen, is hier onderzocht. Er kan nog weinig gezegd worden over verschillen in samenhang tussen tekstniveaus. Om een completer beeld te krijgen van de samenhang van vereenvoudigde teksten, is het wenselijk dat er ook onderzoek wordt gedaan naar de explicitering van referentiële coherentie en het gebruik van woorden die tijd, ruimte en causaliteit uitdrukken.⁹

Voor vervolgonderzoek naar de samenhang van vereenvoudigde teksten, zou gezocht kunnen worden naar materiaal waarbij verschillende niveaus nog sterker inhoudelijk overeenkomen. De Canon van Nederland bevat teksten die over hetzelfde onderwerp gaan en op verschillende niveaus geschreven zijn, maar uit de kwalitatieve analyse is gebleken dat het wel of niet gebruiken van connectieven soms veroorzaakt wordt door het weglaten (of toevoegen) van inhoud. Wellicht is er ander materiaal te vinden waarin de inhoud nog meer overeenkomt tussen niveaus; bijvoorbeeld (herschreven) brieven of folders waarmee mensen geïnformeerd worden. Het zou ook interessant zijn om materiaal te nemen uit een andere context, zoals communicatiemateriaal vanuit de gezondheidszorg of de overheid, zodat bekeken kan worden of in die context dezelfde resultaten naar boven komen.

4.2 Conclusie

Ondanks dat connectieven de laaggeletterde lezer zouden kunnen helpen, worden ze in de teksten van de Canon van Nederland niet extra ingezet op lagere tekstniveaus. Temporele connectieven worden op niveau 1 (B1; kinderen 8-12 jaar en laaggeletterden) en 2 (B2; kinderen 12-15 jaar) wel meer ingezet dan op het niveau 3 (C1; bovenbouw vo en breed publiek). Contrastieve connectieven worden juist op niveau 3 meer ingezet dan op het niveau 2. Echter, bij kwalitatieve analyse van de teksten komt naar voren dat de verschillen mogelijk deels veroorzaakt worden door kleine verschillen in de inhoud van de teksten. Er is meer (kwalitatief) onderzoek nodig om de resultaten van het huidige onderzoek te verifiëren en in context te plaatsen.

⁹ In Pander Maat, Kraf en Dekker (2020) (p. 44) staat meer uitleg over deze tijd-, ruimte- en causaliteitswoorden.

Literatuurlijst

- Buisman, M. & Houtkoop, W. (2014). *Laaggeletterdheid in kaart*. ECBO & Stichting Lezen en Schrijven. Geraadpleegd op 2 maart 2022, van <https://ecbo.nl/onderzoekspublicatie/laaggeletterdheid-in-kaart>.
- Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, 103(2), 429-441.
- Canon van Nederland (z.d.). *Canon van Nederland*. Geraadpleegd op 5-17 april 2022, van <https://www.canonvannederland.nl>.
- Commissie Herijking Canon van Nederland (2020). *Open vensters voor onze tijd: de Canon van Nederland herijkt*. Stichting entoen.nu. Geraadpleegd op 17 maart 2022, van <https://www.canonvannederland.nl/nl/downloads>.
- Crossley, S., Allen, D., & McNamara, D. (2012). Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research*, 16(1), 89-108.
- Crossley, S., & McNamara, D. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In *Proceedings of the Annual Meeting of the Cognitive Science Society* 32(32), 984-989.
- Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and writing*, 15(7), 739-757.
- Den Ridder, J., Josten, E., Boelhouwer, J., & Van Campen, C. (2020). *De sociale staat van Nederland 2020*. Sociaal en Cultureel Planbureau (SCP). Geraadpleegd op 2 maart 2022, van <https://digitaal.scp.nl/ssn2020>.
- Dienst Publiek en Communicatie. (z.d.). *Taalniveau B1*. CommunicatieRijk. Geraadpleegd op 2 maart 2022, van <https://www.communicatierijk.nl/vakkennis/rijkswebsites/aanbevolen-richtlijnen/taalniveau-b1>.
- Evers-Vermeul, J., & van Rooijen, C. (2017). Evidence-based adviezen voor begrijpelijk schrijven, een utopie? *Tijdschrift voor Taalbeheersing*, 39(2), 231-244.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221-233.
- Graesser, A., McNamara, D., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of tekst on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. <https://www.communicatierijk.nl/vakkennis/rijkswebsites/aanbevolen-richtlijnen/taalniveau-b1>.

- Kamalski, J., Sanders, T., & Lentz, L. (2008). Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: Sorting things out. *Discourse Processes*, 45(4-5), 323-345.
- Land, J., Sanders, T., Lentz, L., & Van den Bergh, H. (2002). *Tekstbegrip en tekstwaardering op het vmbo: welke tekstkenmerken dragen bij aan de kwaliteit van studieboekteksten?* Stichting Lezen. Geraadpleegd op 28 februari 2022, van <https://www.lezen.nl/publicatie/tekstbegrip-en-tekstwaardering-op-het-vmbo>.
- Land, J., Sanders, T., & Van den Bergh, H. (2008). Effectieve tekststructuur voor het vmbo: een corpus-analytisch en experimenteel onderzoek naar tekstbegrip en tekstwaardering van vmbo-leerlingen voor studieteksten. *Pedagogische Studiën*, 85(2), 76-94.
- Liu, C., Kemper, S., & Bovaird, J. (2009). Comprehension of health-related written materials by older adults. *Educational Gerontology*, 35(7), 653-668.
- McNamara, D., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22(3), 247-288.
- Noordman, L., & Maes, A. (2000). Het verwerken van tekst. In A. Braet (Red.), *Taalbeheersing als Communicatiewetenschap: Een overzicht van theorievorming, onderzoek en toepassingen* (pp. 29-60). Coutinho.
- Pander Maat, H., & Ditewig, S. (2017). Hoe worden onderwijsteksten vereenvoudigd, en helpt dat? *Tijdschrift voor Taalbeheersing*, 39(2), 245-263.
- Pander Maat, H., Kraf, R., & Dekker, N. (2020). *Handleiding T-Scan*. Geraadpleegd op 23 februari 2022, <https://webservices.cls.ru.nl>.
- Pander Maat, H., Kraf, R., Van den Bosch, A., Dekker, N., Gompel, M., Kleijn, S., Sanders, S., & Sloot, K. (2014). T-Scan: a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands*, 4, 53-74.
- Sanders, T. (2001). Structuursignalen in informerende teksten. *Tijdschrift voor taalbeheersing*, 23(1), 1-21.
- Sanders, T., Land, J., & Mulder, G. (2007). Linguistics markers of coherence improve text comprehension in functional contexts. *Information Design Journal*, 15(3), 219-235.
- Stichting Lezen en Schrijven (2021). *Eenvoudige taal voor laaggeletterden*. Geraadpleegd op 8 mei 2022, van <https://www.lezenenschrijven.nl/wat-doen-wij/oplossing-voor-je-vraagstuk/factsheet-eenvoudige-taal>.

- Van Dooren, W., Van den Bergh, H., & Evers-Vermeul, J. (2012). Leesbare teksten? Over de invloed van structuurmarkeringen op het tekstbegrip en de tekstwaardering van zwakke en sterke lezers. *Levende Talen Tijdschrift*, 13(4), 31-38.
- Van Silfhout, G., Evers-Vermeul, J., Mak, W., & Sanders, T. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*, 106(4), 1036-1048.
- Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2014). Establishing coherence in schoolbook texts: How connectives and layout affect students' text comprehension. *Dutch Journal of Applied Linguistics*, 3(1), 1-29.

Bijlagen

Bijlage A: Connectievenlijst in T-Scan

Causaal

alleen dan	dus	ingevolge	wanneer VG
aangezien	ergo	krachtens	want
anders	ermee	met behulp van	wegens
bijgevolg	erom	middels	zodat
blijkens	ertoe	mits	zodoende
daar	getuige VZ	namelijk	zolang
daardoor	gezien VZ	om VG	dan ook
daarmee	hierdoor	omdat	tengevolge van
daarom	hiermee	opdat	vandaar dat
daartoe	hierom	teneinde	zo VG
daarvoor	hiertoe	vanwege	zo ja
dan ook	hiervoor	vermits	zo nee
dankzij	immers	waardoor	zo niet
derhalve	in verband met	waarmee	zodoende
dientengevolge	indien	waarom	
doordat	ingeval	waartoe	

Comparatief

alsof	meest	naargelang	meer dan
dan VG	minder	naarmate	minder dan
meer	minst	zoals	net zo min

Contrastief

al VG	doch	niettegenstaande	weliswaar
althoewel	echter	niettemin	in plaats
althans	enerzijds	nochtans	laat staan
anderzijds	evengoed	ofschoon	ook al
behalve	evenwel	ondanks	in plaats daarvan
behoudens	hoewel	ongeacht	in tegenstelling tot
daarentegen	hoezeer	tenzij	zonder dat
daarvan	integendeel	terwijl	
desondanks	maar VG	uitgezonderd	

Opsommend (kolom 1: woordgroepniveau; kolom 2-4: zinsniveau)

alsmede	buitendien	ook nog	zelfs
alsook	daarenboven	ook nog eens	zowel
annex	daarnaast	op de eerste plaats	daarbij komt
en	eveneens	op de tweede plaats	dan wel
evenals	evenmin	op de derde plaats	ten eerste
noch	hetzij	op de vierde plaats	ten tweede
of	hierenboven	temeer	ten derde
ofwel	hoofdzakelijk	tevens	ten vierde
respectievelijk	met name	vooral	ten overvloede
zowel	nog eens	voornamelijk	
bovenal	om te beginnen	voorts	
bovendien	ook	waarnaast	

Temporeel

aanstonds
achtereenvolgens
aldoor
aleer
altijd
alvast
alvorens
bijtijds
binnenkort
daarna
daarnet
daarstraks
daartussendoor
dadelijk
destijds
eensklaps
eer VG
eerdaags
eerdat
eerlang
eerst BW
eertijds
eindelijk
ertussendoor
hoelang
indertijd
ineens

ingaaende
inmiddels
meestal
meteen
nadat
naderhand
nadezen
nadien
net BW
olim
onderwijl
onlangs
opeens
pardoes
pas BW
plots
plotsklaps
recentelijk
reeds
sedertdien
sinds
sindsdien
steeds
strakjes
straks
subiet
tegelijk
tegelijkertijd
terstond

tevorens
tezelfdertijd
thans
toen
toenmaals
toentertijd
totdat
vervolgens
vooraf
vooraleer
vooralsnog
voordat
voorheen
wederom
weer BW
weldra
weleer
zodra
zo-even
zo even
zo gauw
zogauw
zojuist
zolang
zonet
zopas
a la minute
hic et nunc

Bijlage B: Verwijderde citaten

Hieronder staan de tekstgedeeltes die niet in de analyse zijn meegenomen. Dit zijn alle (lange) citaten (en bijbehorende vertalingen) die in de teksten voorkomen.¹⁰ Als er in de kopjes hieronder een niveau niet genoemd wordt, betekent dat dat het citaat in de tekst van dat niveau niet opgenomen was.

Tekst 4 (Willibrord) – niveau 2 en 3

“Ec forsacho diabolae, end allum diobolghedae, end ec forsacho allum dioboles wercum and wordum, Thunaer ende Woden ende Saxnote ende allum them unholdum the hira genotas, sint. Ec gelobo in Got alamehtigan fadaer. Ec gelobo in Crist gotes suno. Ec gelobo in Halogan Gast.”

“Ik zweer de duivel af, en alle duivelsoffers, en ik verzaak alle werken en woorden van de duivel, Donar, Wodan, Saxnot en de afgoden die hun gezellen zijn. Ik geloof in God de almachtige Vader, in Christus Gods Zoon en in de Heilige Geest.”

Tekst 6 (Hebban olla vogala) – niveau 1, 2 en 3

“Hebban olla vogala nestas hagunnan hinase hic anda thu, wat unbidan we nu?”
(twee keer weggehaald: 1^e alinea en 2^e alinea)

“Thie wingardon bluoyent anda thee bluom macot suoten stank (‘De wijngaarden bloeien en de bloem verspreidt een zoete geur’).”

Tekst 30 (De grondwet) – niveau 1, 2 en 3

“Allen die zich in Nederland bevinden, worden in gelijke gevallen gelijk behandeld. Discriminatie wegens godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht of op welke grond dan ook, is niet toegestaan.”

Tekst 36 (Anton de Kom) – niveau 2 en 3

“Sranang mijn vaderland. Eenmaal hoop ik u weer te zien. Op den dag waarop alle ellende uit u weggewischt zal zijn.”

Tekst 39 (Indonesië) – niveau 2 en 3

“Proklamasi. Kami bangsa Indonesia dengan ini menjatakan kemerdekaan Indonesia...”

“Wij, het volk van Indonesië, verklaren hierbij dat Indonesië onafhankelijk is...”

¹⁰ In tekst 6 is de vertaling van het citaat onbedoeld niet weggehaald. Dit is bij de andere teksten wel gedaan.